

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 11:08:20 ; Search time 13.29 seconds
(without alignments)
1467.697 Million cell updates/sec

Title: US-09-911-513-2
Perfect score: 2758
Sequence: 1 MKRDHHHHQDKKTMNNEE.....MLGWHTRPLIATSAWKLSTN 532

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	112.5	4.1	772	1 NFL1_HUMAN	Q14494 h nuclear f
2	107	3.9	1158	1 CNDL_SCHPO	Q34679 schizosacch
3	106.5	3.9	1420	1 YMBB_YEAST	Q03496 saccharomyc
4	106	3.8	3206	1 POLG_PSBMV	P29152 p genome po
5	104.5	3.8	953	1 LKAC_PASHA	P55116 pasteurella
6	104.5	3.8	4568	1 DYHC_CAEEL	Q19020 caenorhabdi
7	102.5	3.7	953	1 LKAB_PASHA	P55118 pasteurella
8	101.5	3.7	580	1 RUBA_PASHA	Q42694 chlamydomon
9	101.5	3.7	953	1 LKAL_PASHA	P16535 pasteurella
10	101	3.7	414	1 SPYA_CALJA	P31029 callithrix
11	100	3.6	1007	1 MA2B_FELCA	Q46432 felis silve
12	99.5	3.6	1181	1 YY02_METJA	Q60301 methanococc
13	99.5	3.6	1436	1 DPO3_STAUA	Q53665 staphylococ
14	99	3.6	414	1 SPYA_FELCA	P41689 felis silve
15	99	3.6	636	1 PHBC_XRHET	Q52728 r poly-beta
16	98.5	3.6	558	1 ORC2_XENLA	Q19628 xenopus lae
17	98.5	3.6	4590	1 FATH_HUMAN	Q14517 homo sapien
18	97.5	3.5	1013	1 A60D_DROME	P91927 drosophila
19	96.5	3.5	761	1 PQOF_KLEPN	P27508 klebsiella
20	96.5	3.5	851	1 OBP_HSV11	P10193 herpes simp
21	96	3.5	350	1 SUCA_PIG	Q19069 sus scrofa
22	96	3.5	546	1 PGMO_ECOLI	P36938 escherichia
23	96	3.5	787	1 Y091_CAEEL	P41842 caenorhabdi
24	95	3.4	527	1 TCPB_SCHPO	Q10147 schizosacch
25	95	3.4	665	1 ATKB_TREAC	P57700 thermoplas
26	95	3.4	1549	1 GLSF_CYAC	Q19906 cyanidium c
27	94.5	3.4	614	1 Y156_ARCFU	Q28422 archaeoglob
28	94.5	3.4	1260	1 YAOE_SCHPO	Q10093 schizosacch
29	94.5	3.4	1597	1 CTRO_MOUSE	P49025 mus musculu
30	94	3.4	1071	1 CARB_BACSU	P25994 bacillus su
31	94	3.4	4563	1 APB_HUMAN	P04114 homo sapien
32	93.5	3.4	877	1 DPO1_STRPN	P13252 streptococc
33	93.5	3.4	1330	1 VCAP_PRVIS	Q00705 pseudorabie

RESULT 1

ID	NFL1_HUMAN	STANDARD;	PRT;	772 AA.
AC	Q14494; 012877;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	NUCLEAR FACTOR ERYTHROID 2 RELATED FACTOR 1 (NF-E2 RELATED FACTOR 1)			
DE	(NF-E2-RELATED FACTOR 1) (NUCLEAR FACTOR, ERYTHROID DERIVED 2, LIKE 1)			
DE	(TRANSCRIPTION FACTOR 11) (TRANSCRIPTION FACTOR HBZ17) (TRANSCRIPTION			
DE	FACTOR LCR-FL) (LOCUS CONTROL REGION-FACTOR 1).			
GN	NFE2L1 OR NRE1 OR TCF11 OR HBZ17.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RA	Luna L., Johnsen O., Skartlien A.H., Pedetour F., Turc-Carel C.,			
RA	Prydz H., Kolstoe A.-B.;			
RT	"Molecular cloning of a putative novel human bZIP transcription			
RT	factor on chromosome 17q22.;"			
RL	Genomics 22:553-562(1994).			
RL	[2]			
RP	SEQUENCE OF 326-772 FROM N.A.			
RP	MEDLINE=94310069; PubMed=8036168;			
RA	Caterina J.J., Donze D., Sun C.W., Ciavatta D.J., Townes T.M.;			
RT	"Cloning and functional characterization of LCR-FL: a bZIP			
RT	transcription factor that activates erythroid-specific, human globin			
RT	gene expression.;"			
RL	Nucleic Acids Res. 22:2383-2391(1994).			
CC	-!- FUNCTION: ACTIVATES ERYTHROID-SPECIFIC, GLOBIN GENE EXPRESSION.			
CC	-!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).			
CC	-!- SIMILARITY: BELONGS TO THE BZIP FAMILY. CMC SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed, usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X77366; CAA54555.1; -			
DR	EMBL; U08853; AAA20466.1; -			
DR	HSP; P34707; LSKN.			
DR	MIM; 163260; -			
DR	InterPro; IPR001871; bZIP.			
DR	Pfam; PF00170; bZIP; 1.			
DR	SMART; SM00338; BRZL; 1.			
DR	PROSITE; PS00036; bZIP_BASIC; 1.			
KW	Transcription regulation; Activator; DNA-binding; Nuclear protein.			
FT	DOMAIN 125 288 ASP/GLU-RICH (ACIDIC).			
FT	DNA_BIND 496 517 POLY-SER.			
FT	DNA_BIND 659 674 BASIC MOTIF.			

ALIGNMENTS

Q64535 rattus norv
P31030 oryctolagus
O04015 a delta l-p
O84303 chlamydia t
Q43933 homo sapien
P07210 human rhino
P26595 mus caroli
P00561 escherichia
P53254 saccharomyc
P10955 rhizobium m
Q29052 sus scrofa
O13776 schizosacch

34 93.5 3.4 1451 1 AT7B_RAT
35 93 3.4 392 1 SPYA_RABIT
36 93 3.4 717 1 P5CS_ACTCH
37 93 3.4 934 1 PKNI_CHLTR
38 93 3.4 1283 1 PEXI_HUMAN
39 93 3.4 2164 1 POLG_HRV89
40 92.5 3.4 412 1 A1AT_MUSCR
41 92.5 3.4 820 1 AK1H_ECOLI
42 92.5 3.4 1237 1 YG2L_YEAST
43 92 3.3 505 1 FIXL_RHIME
44 92 3.3 902 1 ITHL_PIG
45 92 3.3 925 1 YE9G_SCHPO

```
FT DOMAIN 582 704 LEUCINE-ZIPPER.
SQ SEQUENCE 772 AA; 84703 MW; C868807C6046BEF5 CRC64;

Query Match
Best Local Similarity 4.18; Score 112.5; DB 1; Length 772;
Matches 118; Conservative 64; Mismatches 181; Indels 205; Gaps 27;

QY 55 MMSVQEDDLSQLATETVHYNPAELYTHVMDLTPSSNAEYDLKAIPGDAILNOFA 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 83 LLSQVRLDRQVPT-----TEVNAW-----LVHRDEGS-----VSGSQNSGLA 123

QY 115 IDSASSNQGGGGTYTNKRLKCSNGVETTTATSTRVHLVDSOENGRLVHALLA 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 124 LESSGLQDVTPD-----NGVRETEQGFGE-----DLEDLG-----A 158

QY 175 CAEAVQ-----KENTVBAELVKQIGFLAVSOIGAMRKVATVF----- 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 159 VAPPVSGDLTKEDIDLILWRQ-----DIDLGAGREVFDYSHRQKQDVKEKLRDGEQ 213

QY 213 -----AEALARRIY-----RLSPQSPIDH-----SLSDTLQMHFYETCPYIKFAHF 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 214 DTWAGEGAELARNLLVDGETGESFPAGVPSGEDQTALSLECLRL-LEATCPGENAEF 272

QY 255 TAN-QAILEAQKRVHVID-----FMSQGLQWLPALMQALALR----- 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 273 PADISSITEAVPSEPPALQNNLLSPLLTGCTESPFDLEQ--QWQDLMSIMEMQAMEVNT 330

QY 294 -----PGGPVRLTGTGPAPDNFDY-LHEV---GCKLAHLAEAIHVEFEYRGFV 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 331 SASEILYSAPGDPPLSTNYSIAPNTPINQVSLHQASLGCGSQDFLLFSPEVE----- 383

QY 341 ANTLDLADLAEMLERLPSIESVAVNSVF-----ELHKLGLR 376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 384 -SLPVASSSTLLPLAPS--NSTSLNSTFGSTNLGTGLFPPPLNGTANDTAGPELPDPLG- 439

QY 377 PGADKV-----LGVVNQIKPEFTVVEQSNHNSPIFLDRFTESLHYSTLPDSL 427
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 440 -GLIDEAMLDEISLMDLAEIEGFPVQASQLEEFDSQSLSD-----SSHPSL-SSS 493

QY 428 EGVPSGQDKVMSEVYLQKQICNVVACDGPDRVERHETLSQWRNRFSGAGFAAHIGSNAF 487
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 494 EGSSSSS-----SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 526

QY 488 QKASMLIAL--FNGGEGYRVESDGCLM 513
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 527 SSDSETLDLEAEAGVGYQPEYSKFCRM 554

RESULT 2
CNDL_SCHPO STANDARD; PRT; 1158 AA.
AC 094679;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CONDENSIN COMPLEX SUBUNIT 1 (P128).
GN CND1 OR SPBC776.13.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 518-536; 658-667 AND 1126-1138.
RX MEDLINE=99415811; PubMed=10485849;
RA Sutani T., Yuasa T., Tomonaga T., Dohmae N., Takio K., Yanagida M.;
RT "Fission yeast condensin complex: essential roles of non-SMC subunits
RL for condensation and Cdc2 phosphorylation of Cut3/SMC4."
RN Genes Dev. 13:2271-2283(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;

```

```
RA Lyne M., Rajandream M.A., Barrell B.G., Wedler H., Kutzner M.,
RL Wambutt R.;
RA Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR CHROMOSOME CONDENSATION IN MITOSIS.
CC ESSENTIAL FOR VIABILITY.
CC -!- SUBUNIT: PART OF THE CONDENSIN COMPLEX.
CC -!- SUBCELLULAR LOCATION: NUCLEAR, AND CYTOPLASMIC (DURING
CC INTERPHASE).
CC -!- SIMILARITY: BELONGS TO THE CONDENSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB030212; BAA82624.1; -.
CC EMBL; AL035263; CAA22886.1; -.
CC Mitosis; Nuclear protein.
CC SEQUENCE 1158 AA; 131323 MW; 37A7A2206830DF3D CRC64;

Query Match 3.98; Score 107; DB 1; Length 1158;
Best Local Similarity 19.28; Pred. No. 7.7;
Matches 100; Conservative 69; Mismatches 155; Indels 198; Gaps 23;

QY 22 DGNMGDELLAVLGYKVRSEMAVQAQLEQLEVMSNVQEDDLSQLATETVHYNPAELYT 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 53 DEDCLEKLFAL-----CSHFAD-----LSSVNRKVDLLTSNSESAILSD 95

QY 82 WLDSMLTDLNPPSSNAEYDLKAIPGDAILNQFAIDSASSNQGGGGTYTNKRLKCSNG 141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 96 MISANATDFTVPQNLN-----TTGIAFQLTVNSLSSNQ-----LS 132

QY 142 VVETTTATSTRVHLVDSOENGRLVHALL-ACAFAVQKENTVBAELVKQIGFLAVS 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 133 VIRSTNTVGRKKNPTNTNNGISHVNALDAILTLFQKK-----LS 176

QY 201 QIGAMRKVATVFAEALARRIYRLSPSQSPIDHS-----LSDTLQMHFEYETCPYIKF 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 177 RVWTTSSDERDMFLSLFKPKPIYTLMESEINIKNASFRSLFNIGLAVQFH----- 226

QY 252 AHFTANQAILAEFOGKRVHVIDFMSQGLQWLPALMQALALRPGGPVRLTGTGPPAPD 311
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 227 NHTTAAET-----NIQNLQY----- 242

QY 312 NFDYLHEVGCKLAHLAEAIHVEFE-----YRGFVANTLADLD-----ASMLELR 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 243 -FEHLSEYAADLVHI---VTQFNSVTLAEGII-RTLCGLEFNNDNDVKGPQKQVALFLVRL 297

QY 356 PSEIESVAVNSVFELHKLGLRPG-----AIDKVLG--VVNQIKPEFTVVEQSNHNSP- 407
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 298 SSLIPNCLAKQLTQLVLLDSESYTLRCAIIEVLAVNVVIDQIHDE-----AQENSESVPA 353

QY 408 -----IFLDRF-----TESLHYSTLTFLSLEGVPSGQDKVMSEVYLQKQICNVVA 452
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 354 TVQSLMDLLSERLLDISPCYRTKVLHVFIKIFDLPKPYPRKRQEI-AEL-----VIR 404

QY 453 CDGPDVRVERHETLSQWRNRFSGAGFAAHIGSNAPKQASMLL 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 405 C-LQDR-----SSHVRNRNAIKLFSKLL 425

RESULT 3
YMBL YEAST
ID YMBL YEAST STANDARD; PRT; 1420 AA.
AC Q03496; Q04862;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 163.6 KDA PROTEIN IN PET111-TIF11 INTERGENIC REGION.

```


DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00863; Peptidase_C4; 1.
DR Pfam; PF00851; Peptidase_C6; 1.
DR Pfam; PF00767; Poty_coat; 1.
DR Pfam; PF01577; Poty_P1; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR PRINTS; PR00966; NIAPOTIPTASE.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KW ATP-binding.
FT CHAIN 1 ? N-TERMINAL PROTEIN.
FT CHAIN ? 856 HELPER COMPONENT PROTEINASE.
FT CHAIN 857 ? PROTEIN P3.
FT CHAIN ? 1266 6 KDA PROTEIN 1.
FT CHAIN 1267 1902 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1903 1955 6 KDA PROTEIN 2.
FT CHAIN 1956 ? GENOME-LINKED PROTEIN.
FT CHAIN ? 2395 NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2396 2915 NUCLEAR INCLUSION PROTEIN B.
FT CHAIN 2916 3206 COAT PROTEIN.
FT SITE 1266 1267 CLEAVAGE (BY 49 KDA PROTEASE).
FT SITE 1902 1903 CLEAVAGE (BY 49 KDA PROTEASE).
FT SITE 1955 1956 CLEAVAGE (BY 49 KDA PROTEASE).
FT SITE 2395 2396 CLEAVAGE (BY 49 KDA PROTEASE).
FT SITE 2915 2916 CLEAVAGE (BY 49 KDA PROTEASE).
FT BINDING 2016 2016 COVALENT LINKAGE OF VIRAL RNA (BY SIMILARITY).
FT NP_BIND 1351 1358 ATP (POTENTIAL).
FT SEQUENCE 3206 AA; 364271 MW; 42A3D921B9A0CBF CRC64;
Query Match 3.8%; Score 106; DB 1; Length 3206;
Best Local Similarity 20.08; Pred. No. 40;
Matches 120; Conservative 84; Mismatches 170; Indels 226; Gaps 33;
Qy 8 HQDKKMM--NEEDDNGM-DELLAVLYK-----VRSEMAADV 45
Db 70 HLVDKTPKRGSETEGDGMAEAIAKVTGAEPVNCFMVGTICKINENSVKGVMAAI 129
Qy 46 AOKLEQVVMNSVQEDDLSLATETHYNPAE---LYTWLDSMLTDLNPPSSNAEYDLK 102
Db 130 PRQITQDEVFMRKARLQ--AAVASTIEREKEKQFAESKLEELRARREKLKG-IVIK 186
Qy 103 AIPG-----DAILNFAIDSASSSSGOGGDT-----YTTNKRKLCSSNGVVETTTA---- 148
Db 187 TRKGLEWREATPNQGRKLQSTFSDASGGKTLTPHTIYCKTKSKFSNGVKCATSKMR 246
Qy 149 -----TAESTRHVVLDVDSQ-----NGVRLV----- 169
Db 247 TVRKPSLKMKTESIDVLIEQVMTIAGKHAKOVTLIDKQKTRNVWVRVNGVRLQLVETK 306
Qy 170 -HALLACAEAVQKENTLVAEALVKQIGFVLSQIGAMRKVATYFAEALRRIYRLSPQS 228
Db 307 HKGIISOKDASLNLIT-----KRVARHF-----ARKTAYIHPSDS 342
Qy 229 PIDHSLSDTLQMHYETCPYLYKFAHFTANQAI-----LEAFQK-----KRV 270
Db 343 -ITHGSGV-----FLR-ANISGKSYSIDDLFVVRGKRNGKLMSRNKVAWRKM 391
Qy 271 HVIDFSMSOGLQ-WPAL-MOALALRP-----GG-----PPVPERL 302
Db 392 FOIDHFSIVGIRKINWAFDAEVYKURDESVDSDHDCVGGITPECGILAAQILRVFPCWRI 451
Qy 303 T-----GIGPPAPDNFDYHLEYGCKLA-----HLAEAIHVEFEYRGFVANTLADLA 349
Db 452 TCTKCISNWSLSPQSEQIEHIIYERG-NLAIQDLNKRIPSAHHV-----T 494
Qy 350 SMLELRSETESVAVNSVFEI-----HKLGRPGCATDKVLGVVQNQKPETF-----TVV 398
Db 495 QMVELLQRIRK-----NTTFDMGNNTKVHELIGH--RODGVPRHLNRLNLSLAANGSSTI 548

Qy 399 EQESNHNSPIFLDRP-----TESLHYSTLFDLSLEGVPSGQDKVMSEVYLGKQICNVVACD 454
Db 549 EWESNMSELLELARWHNKRTESI-----ASGGISSFRNKRISAKA-----QINFALMCD 596
RESULT 5
LKA3_PASHA STANDARD; PRT; 953 AA.
AC P55116;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LEUKOTOXIN FROM SEROTYPE T3.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Mannheimia.
OX NCBI_TaxID=75985;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE T3;
RC MEDLINE=94041617; PubMed=8225575;
RA Burrows L.L., Lo R.V.C., Olah-Winfield E.;
RT "Molecular analysis of the leukotoxin determinants from Pasteurella
haemolytica serotypes 1 to 16";
RL Infect. Immun. 61:5001-5007(1993).
CC -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
DEFINED.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
ACTIVITY.
CC -!- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC -!- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN
MODIFIED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; U01216; AAB36691.1; -.
DR HSSP; P02392; ICTF.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003353; RTX_N.
DR Pfam; PF00353; hemolysinCabin; 1.
DR Pfam; PF02382; RTX; 1.
DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
DR Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;
KW Transmembrane; Lipoprotein; Palmitate.
KW TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 297 318 POTENTIAL.
FT TRANSMEM 381 401 POTENTIAL.
FT DOMAIN 734 784 6 X REPEATS, GLY-RICH.
FT REPEAT 734 739 1.
FT REPEAT 743 748 2.
FT REPEAT 752 757 3.
FT REPEAT 761 766 4.
FT REPEAT 770 775 5.
FT REPEAT 779 784 6.
SQ SEQUENCE 953 AA; 101948 MW; FDBDCE2FDC85FDF2 CRC64;

Query Match 3.8%; Score 104.5; DB 1; Length 953;
Best Local Similarity 18.8%; Pred. No. 8.8;
Matches 124; Conservative 97; Mismatches 199; Indels 239; Gaps 33;

[illegible]


```

DR Pfam: PF00118; cpn60_TCP1; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONINS.CPN60; 1.
KW Chaperone; ATP-binding; Chloroplast; Transit peptide; Heat shock.
FT TRANSIT 1 580 CHLOROPLAST (POTENTIAL).
FT CHAIN ? 580 RUBISCO SUBUNIT BINDING-PROTEIN ALPHA
FT SUBUNIT.
SQ SEQUENCE 580 AA; 61863 MW; 16FD34B115E706F7 CRC64;

Query Match 3.7%; Score 101.5; DB 1; Length 580;
Best Local Similarity 21.7%; Pred. No. 7.1;
Matches 100; Conservative 45; Mismatches 129; Indels 187; Gaps 24;

QY 65 SOLAETVTHYPAELYTWLDSMLTNPSPS-----SNAEY-----DLKAIPG-DAIL 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 126 SVLAEMTHYG-----LQSVTAGANPIAVKRGDKTAELVAKLKEHAKPVKGRDDIK 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 111 NOFAIDSASSNOGGGGDTYTNKRLKCSNGV--VETTTAT-----
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 179 N---VASISAGNDNAIGEMADALDKVGSNGVLSIETSNSTETVVVEQGEIDRGYISP 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 150 -----AESTRHVVLVDSQ-----ENGRLVHALLACAEAVQKLENL--- 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 QFVTNQERLLVEYDNCRLVLTQDKIDAIRDIPILEQVTRLNAPLLIITAEDVSGEALATL 295
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 185 -----TVAEALVKOIGF-----LAVSOIGAMRK 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 296 VVNLKRGVLNVCATKAPGFGERRKSLLODIAIVTGAEFIAKDGLMKVQAVVEQLGVARK 355
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 208 V-----ATYFAEALAR-----RIYRLSPQSPIDHSLSDTLQMHFETCPYLKFAHPTA 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 356 VTVANNTTLLIADAASKDEIEMRIATLKKELAETD-SYVDTEKLS-----ERIKLSG 407
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 257 NQAIL-----EAFQKKRVHVVIDFSMSQGLQWPAQMQLALRPGGPPVFLRTGIGPPAP 310
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 408 GVAVIKVGAANYAELEDEKRLIED---AKNATFAAVEEGIV--PGG----- 448
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 311 DNFDYLHEVGCKLAHLAEAIHVEFEYRGFVANTLAD-----LDASMLELR-PSEI--ES 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 -----GAALLHLSLVPAFKE-----TLTDAEEKLGADIVMKSLRACPRLIADN 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 362 VAVNSVFELHKLGRP-----GAIDKV-----LGVVNQIK 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 493 AGVEGEIVQRLGKPFEGVGYNAMIDKVENLDDAGVIDPAK 533
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
ID LKAL_PASHA STANDARD; PRT; 953 AA.
AC P16535;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 30-MAR-2000 (Rel. 39, Last annotation update)
DE LEUKOTOXIN FROM SEROTYPE A1.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE A1;
RX MEDLINE=87306837; PubMed=3040588;
RA Lo R.Y.C., Strathdee C.A., Shewen P.E.;
RT "Nucleotide sequence of the leukotoxin genes of Pasteurella
   haemolytica A1."
RL Infect. Immun. 55:1987-1996(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE A1 / PHL101;
RX MEDLINE=89210283; PubMed=2707120;

```

```

RA Highlander S.K., Chidambaram M., Engler M.J., Weinstock G.M.;
RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene
   cluster."
RL DNA 8:15-28(1989).
RN [3]
RP SEQUENCE OF 884-953 FROM N.A.
RC STRAIN=SEROTYPE A1 / PHL101;
RX MEDLINE=90236888; PubMed=2185213;
RA Highlander S.K., Engler M.J., Weinstock G.M.;
RT "Secretion and expression of the Pasteurella haemolytica Leukotoxin."
RL J. Bacteriol. 172:2343-2350(1990).
CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
   CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
   DEFINED.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
   CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
   ACTIVITY.
CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
   INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC -1- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN
   MODIFIED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (see http://www.isb-sib.ch/announce/
   or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M20730; AAA25529.1; -
DR EMBL: M24197; AAA25543.1; -
DR PIR: S29516; S29516.
DR HSPR: P02392; ICTF.
DR InterPro: IPR001343; Hemlysn_Ca_bind.
DR InterPro: IPR003355; RTX_N.
DR Pfam: PF00353; hemolysinCabin; 1.
DR Pfam: PF02382; RTX; 1.
DR PRINTS: PR00313; CABNONGRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCIUM; 4.
KW Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;
KW Transmembrane; Lipoprotein; Palmitate.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 297 318 POTENTIAL.
FT TRANSMEM 366 390 POTENTIAL.
FT DOMAIN 734 784 6 X REPEATS, GLY-RICH (BY SIMILARITY).
FT REPEAT 734 739 1.
FT REPEAT 743 748 2.
FT REPEAT 752 757 3.
FT REPEAT 761 766 4.
FT REPEAT 770 775 5.
FT REPEAT 779 784 6.
FT CONFLICT 409 414 FEHVN -> LSTLQI (IN REF. 2).
FT CONFLICT 742 742 D -> Y (IN REF. 2).
SQ SEQUENCE 953 AA; 101996 MW; 7F93D113A118C05F CRC64;

Query Match 3.7%; Score 101.5; DB 1; Length 953;
Best Local Similarity 18.8%; Pred. No. 15;
Matches 122; Conservative 100; Mismatches 207; Indels 221; Gaps 32;

QY 19 EEDDNGMDLAV---LGYKVRSSSEMDVAQK-----LEQLEVMKSNVQEDD 63
   : : ||| : : || : : || : : || : : || : : || : : || : : || : : || : : ||
DB 55 DTGGNQLDLVKAEEELGIEVQREERNNTATQTSIGTTQTAIGLTERGIVLSAPQIDK 114
   : : ||| : : || : : || : : || : : || : : || : : || : : || : : || : : ||

QY 64 LSQI-----ATETVHNPALYTWLDSMLTNPSPSNAEYDLKAIPGDAILNOFAI 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 115 LLQTKAGQALGSAESIVQNAKAKTVLSIQISGLSVLAGMDLD-EALQNNNS--NQHAL 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 116 DSASSNQGGGDTYTTNKRKCSNGVYETTT-----A 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
Db 172 AKAG-----LELTNSLIENTANSVKTIDFEGEQISQSGSKLQNTKIGLT 215
Qy 149 TAESTRHVVLDQSGVRLVHALLACAEAVOKENLTVAE-----ALVKQIGF---LAVSQI 202
Db 216 LGDKLNKNGGLDKAGLDVLSGLSGATAA-----LVLDKNASTAKVKVAGFELANQVV 271
Qy 203 GAMRK-VATYFAEALARIYRLSPQSQSDISLDTLQMHFYETCPYL-----KFAHFT 255
Db 272 GNITKAVSSYI---LAQRAAGLSSTGPVAALIASTVSL---AISPLAFAGTADKFNHAK 325
Qy 256 ANOATLEAF-----QGKKRVHVIDFSMSOGLQWLPALMOALALRPGG----- 296
Db 336 SLESTAEKFKLGYDGDNLLAEYQGTGTIDASVI-----AINTALAAIAGGVSAAAAG 379
Qy 297 -----PPVRLTGTGPPADPNFYDLYHEVGCKLAHLAEAIH-----VEFE-----YRGF 339
Db 380 SVIASPIALLVSGITGVISTILQYSKQ--AMFEHVANKIHKNIVEWKNHNGKNYFENG 437
Qy 340 VANTLADLDASH---LELRPSEIESVAVNSVFE-----LHKLGRPCAIDKVLGVVNQI 390
Db 438 DARYLANLDQNMKFLNLN-KELQAEYIAITQQQWDMNIGDLAGISRLGEKVL----- 490
Qy 391 KPEIFTVVEQSNHNSPIFLDRFETSLHYVS---TLRPSLEGV---PSGODKVMSEVYL 443
Db 491 -----SGKAYVDFAFEGKHKIKADKLVDLSANGIIDVNSNGRAKTHILFR 536
Qy 444 GKQICNVACDPRVERHET-----LSQWRNRFGSA-----GF 477
Db 537 -----TPLLTPCTEHRVQTKVEYITKLNIRVDSKKITDGAASFTDLTNVQRIGI 591
Qy 478 AAHIGTS-NAPKQASMLAL-----FNGGEGY-RVEESDG 510
Db 592 ELDNAGNVTKTKETKIIAKLGEQDNDVFGSGTTEIDGEGGYDRVHYSG 641

RESULT 10
SPYA_CALJA
ID SPTA_CALJA STANDARD; PRT; 414 AA.
AC P31029;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SERINE-PYRUVATE AMINOTRANSFERASE, MITOCHONDRIAL PRECURSOR
DE (EC 2.6.1.51) (SPT) (ALANINE-GLYOXYLATE AMINOTRANSFERASE)
DE (EC 2.6.1.44) (AGT).
GN AGT OR AGT1.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callithrix.9483;
OX NCBI_TaxID=9483;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339467; PubMed=1339350;
RA Purdie P.E., Lumb M.J., Danpure C.J.;
RT "Molecular evolution of alanine/glyoxylate aminotransferase 1
RT intracellular targeting. Analysis of the marmoset and rabbit genes.";
RL Eur. J. Biochem. 207:757-766(1992).
CC -!- FUNCTION: DUAL METABOLIC ROLES OF GLUCONEOGENESIS (IN THE
CC MITOCHONDRIA) AND GLYOXYLATE DETOXIFICATION (IN THE PEROXISOMES).
CC -!- CATALYTIC ACTIVITY: L-SERINE + PYRUVATE = 3-HYDROXYPYRUVATE +
CC L-ALANINE.
CC -!- CATALYTIC ACTIVITY: L-ALANINE + GLYOXYLATE = PYRUVATE + GLYCINE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX AND PEROXISOMAL.
CC -!- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M84414; AAA35397.1; -.
DR PIR: S24154; S24154.
DR InterPro: IPR000192; AminoTransf_class_v.
DR Pfam: PF00266; aminotran_5; 1.
DR PROSITE: PS00595; AA_TRANSFER_CLASS_5; 1.
KW TRANSFERASE; Aminotransferase; Pyridoxal phosphate; Peroxisome;
KW Mitochondrion; Transit peptide; Alternative initiation.
FT TRANSIT 1 23 MITOCHONDRION
FT CHAIN 24 414 SERINE-PYRUVATE AMINOTRANSFERASE,
FT MITOCHONDRIAL ISOFORM.
FT CHAIN 23 414 SERINE-PYRUVATE AMINOTRANSFERASE,
FT PEROXISOMAL ISOFORM.
FT INT_MET 23 23 FOR PEROXISOMAL ISOFORM.
FT BINDING 231 231 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 414 AA; 45054 MW; 604866DA42EEDDE1 CRC64;

Query Match 3.7%; Score 101; DB 1; Length 414;
Best Local Similarity 21.9%; Pred. No. 4.7;
Matches 70; Conservative 41; Mismatches 115; Indels 94; Gaps 15;

Qy 292 LRPGPPVFRUTGTGPPADPN-----FDYLHEVGCKLA-HLAEAIHVEFYR 337
Db 37 LKPLSIPTRLLLG---PGSNLPPTMAAGGLQMLGHMHKETYQIMDEIKEGIQVFTFR 93
Qy 338 GFVANTLAD-----LDASMLE-LRPSEIESVAVNSVF-----ELHKLGRPG 378
Db 94 NPLTLVLSGSGHCALEALINVLPGSFLGVNGINGQRAADIGERIGARVHPMTKDPG 153
Qy 379 ---AIDKVLGVVNVQIKPEIFTVVEQSNHNSPIFLDRFETSLHYYSTLF-----DSLEGV 430
Db 154 GHYTLQVEVEGLAQHKPVLLFLTHGESSGVLQPLDGLGELCHRYKCLLLVDSVASLGA 213
Qy 431 P-----SGQDKVM-----SEVYLG-KQICNV 451
Db 214 PLYMDQGGIDILYSGSQKVLNAPPGTSLLSFSDTAKNKIYSRKTKPSSFYLDVLYLANLW 273
Qy 452 ACDGPDPRVERHET---LSQWRNRFGSAGFAAAHIGSNFAKQASMLALFNG---GEGYRVE 506
Db 274 GCDGQPMYHHTTPVSVLSYSLREGALLSSEQL-ENSWRKHREAAAYLHGRLOALGLRLF 332
Qy 507 ESDGCLMLGWHTRPLIATSA 526
Db 333 VKDPALRL-----PTVTVA 347

RESULT 11
MA2B_FELCA
ID MA2B_FELCA STANDARD; PRT; 1007 AA.
AC O46432;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LYSOSOMAL ALPHA-MANNOSIDASE PRECURSOR (EC 3.2.1.24) (MANNOSIDASE,
DE ALPHA B) (LYSOSOMAL ACID ALPHA-MANNOSIDASE) (LAMAN).
GN MAN2B1 OR MANB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98060744; PubMed=9396732;
RA Berg T., Tollersrud O.K., Walkley S.U., Siegel D., Nilssen O.;
RT "Purification of feline lysosomal alpha-mannosidase, determination of
RT its cDNA sequence and identification of a mutation causing alpha-
RT mannosidosis in Persian cats.";
RL Blochem. J. 328:863-870(1997).
CC -!- FUNCTION: NECESSARY FOR THE CATABOLISM OF N-LINKED CARBOHYDRATES
CC RELEASED DURING GLYCOPROTEIN TURNOVER (BY SIMILARITY).
```


CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING
CC ALPHA-D-MANNOSE RESIDUES IN ALPHA-D-MANNOSIDES.
CC -1- SUBCELLULAR LOCATION: LYOSOMAL.
CC -1- PTM: PROCESSED INTO 3 PEPTIDES OF 72 KDA, 41 KDA AND 12 KDA.
CC -1- DISEASE: DEFECTS IN MANB ARE THE CAUSE OF LYOSOMAL ALPHA-
CC MANNOSIDOSIS, A LYOSOMAL STORAGE DISEASE CHARACTERIZED BY
CC ACCUMULATION OF UNBRANCHED OLIGOSACCHARIDES CHAINS.
CC -1- SIMILARITY: BELONGS TO FAMILY 38 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF010191; AAB97672.1; -
CC EMBL; AF010192; AAB97733.1; -
CC InterPro; IPR000602; Glyco_hydro_38.
CC Pfam; PF01074; Glyco_hydro_38; 1.
CC Hydrolase; Glycosidase; Glycoprotein; Lyosome; Zymogen; Signal.
CC SIGNAL 1 50 BY SIMILARITY.
CC CHAIN 51 1007 LYOSOMAL ALPHA-MANNOSIDASE.
CC SEQUENCE 1007 AA; 113230 MW; C064A816BF5DC20A CRC64;
CC -----

		Query Match	3.6%; Score 100; DB 1; Length 1007;	
		Best Local Similarity	19.0%; Pred. No. 20;	
		Matches 104; Conservative	80; Mismatches 185; Indels 178; Gaps	32;
QY	74	YNPAELYTWLDSM-----LTDLNPPSSNAFYDLKAIPGDAILNQFAIDSNASSNSQGGG	127	
DB	262	YNPEKLCW-DTLCADKPVEDRRSPYNAE-----ELVNYF-LQLATAO-----G	305	
QY	128	DYTTNKRKLCSNGVVEETTAT--AESTRHVVLDSQE--NG--VRLVHALLAC-AEAV	179	
DB	306	QHFTNTHTMTSGDFOYENANMWFRNLDRLIQLVNAQQANGSRVNVLSTPACYLWEL	365	
QY	180	OKENLTVAEALVKQIGFLAVSQIGAMRKVATYFAEALARIRYLRLSPQSPIDHLSLDTLQ	239	
DB	366	NKANLTWS---VKODDFPYAD-GPHQFWGYESSRPALKRY-----ERLS	407	
QY	240	MHFVETC-----PYLKFAHFTANOAIL-----EAFQGKKRVHVI-DFSMSQ	279	
DB	408	YNFLOVCNQLLEALAGPAANVCPGSGDSAPLNQMAVLQHHDAVSGTSKHIVADDIARQL	467	
QY	280	GLOWP-----ALMOQALRPGGPVPVIRTGITGPPAPDNFDLYLHEVG----CKLAHLAE----	328	
DB	468	AAGWDPCEVLLSALA-----RLSG-----SKEDFTYCRLNVSVCPQLSQTAKNFQV	514	
QY	329	-----ATHVEFEYRGFVA---NTLADLDASMLE-----LRPSEIESAVN	365	
DB	515	TIYNPLGRKIDMWVRLPVSKHGFPVRDPNGTVPSDVVILPSSDGQELLFPASFVALGF-	573	
QY	366	SVELHKLLG-RPCA-----IDKVLGVNQ-----IKPFITVWEQESNNHSPIF	409	
DB	574	SIYSVSQVGPORHAHKPQRSQPWRSKVLAIONEHIRARFPDPTGLLVEMENDQNLL	633	
QY	410	LDRFETSLHY-----STLFDSLEGV-----PSGODKVWSEVYLKQICNVACDGPDRVE	460	
DB	634	PVR-QAFYWNVASVGNLSTQVSGAIFRPNQEKPLMVSHWAQTRLVKT-----PLVQE	686	
QY	461	RHETLSQRNRFGSAFAAAHIGSNAPFKQAQSMILLALENGGEVRYEESDGCMLMGWHTRP	520	
DB	687	VHQNFSAW-----CSQVVRLYRGQRHLELE-----WTVGP	716	
QY	521	LIATS AW 527		
DB	717	IPVGDGW 723		

RESULT 12

YY02_METJA	ID	YY02_METJA	STANDARD;	PRT; 1181 AA.
	AC	Q60301;		
DT	01-NOV-1997	(Rel. 35, Created)		
DT	01-NOV-1997	(Rel. 35, Last sequence update)		
DT	20-AUG-2001	(Rel. 40, Last annotation update)		
DE	HYPOTHETICAL PROTEIN MJEC502.			
GN	MJEC502.			
OS	Methanococcus jannaschii.			
OC	Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;			
OC	Methanococcus			
OX	NCBI_TaxID=2190;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=JAL-1 / DSM 2661 / ATCC 43067;			
RX	MEDLINE=96337999; PubMed=8688087;			
RA	Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,			
RA	Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,			
RA	Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,			
RA	Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,			
RA	Scott J.L., Geoghegan N.S.M., Weldman J.F., Fuhmann J.L., Nguyen D.,			
RA	Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,			
RA	Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,			
RA	Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;			
RT	"Complete genome sequence of the methanogenic archaeon, Methanococcus			
RT	jannaschii.";			
RL	Science 273:1058-1073(1996).			
CC	-1- SIMILARITY: WEAK IN THE C-TERMINUS, TO ADENINE-SPECIFIC			
CC	METHYLTRANSFERASES.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L77119; AAC37060.1; .-			
DR	HSPG; P14385; IAQJ.			
DR	TIGR; MJEC502; .-			
DR	InterPro; IPR002296; NI2NG_mtfrase.			
DR	InterPro; IPR002052; NS_Mtase.			
DR	PRINTS; PR00507; NI2NGMTFRASE.			
DR	PROSITE; PS00092; N6_MTASE.1.			
KW	Hypothetical protein; Transmembrane; Complete proteome.			
FT	TRANSEM 837 857 POTENTIAL.			
SEQ	SEQUENCE 1181 AA; 138438 MW; 9f961d8f8c6a4bdd CRC64;			
Query Match		3.6%;	Score 99.5;	DB 1; Length 1181;
Best Local Similarity		20.3%;	Pred. No. 28;	
Matches 112;		Conservative	85;	Mismatches 199; Indels 157; Gaps 29;
QY	24 NGMDLLAV----	LGKYKVRSSMADVAOKLEQLFYMNSVQEDDLSQLATEFVHYNPAPLY 80		
	:	: : : : : : : : : : : : : : :		
Db	524 SALKELQIKRIYYLLR--EEMDIYK--EKLGILLNNLYGVDDIIAIVEL-----AKLR 574			
QY	81 TWLDSMLTDLNPPSSNAEYDLKAIP-GDAILNQFAIDSASSNSGGGDYTTNKRLKCS 139			
	: : : : : : : : : : : : : : :			
Db	575 LWL-ALTENL-----DVEALKRGVLPLNTIYN-----VRCG 605			
QY	140 NGVVETTATAESTRHVVDVSQENGVRHLVHALACAEVQKNLTVA-EALVKQIGFLA 198			
	: : : : : : : : : : : : : : :			
Db	606 NSLVGWIDENLKQSISYLCNDVMIMCVLEGILIINHNSEERKKLKKAKELLEARDGYVL 665			
QY	199 VSQIGAMRKVATYPAEALARIRYRLSPS-QSPIDHSLSDTLQMHFYTCTPYLKEAFH- 254			
	: : : : : : : : : : : : : : :			
Db	666 DNYVEAYH-----LLYEVIYTSCLKANLLKELLDELIRDSIYESVTYPAYFAEIYQN 716			
QY	255 -----TANQAILEAFQGKKRVH-VIDFSMSOGLQPALMQALALRPGGPVPERL 302			
Db	717 GNKKNGKSKKWRPRVEFEKLPKPHWKIDFG-----WIKEGFDFVLIIGNPPYGNL 770			

```
QY 303 TGIGPPAD-----NFDYLHEVGCKLAHLAEAIHVEFEYRGV-----ANT 343
D 771 --LSPTKEIMKRRDTPEDIEVTF---IVHSSKLLKNE-GYLGFIIIPSSFGTGVRYSNL 824
QY 344 LADLDASM---LELRPSEIESVAV--NSVFLHKLGRPGAKDVL-----GVVN 388
D 825 RKELETFMKCLKLIYLPDFVFGAYVDCIILHK--RPPKSEDILVIYAPFKTKKISF 882
QY 389 QIKPEIFTWVEQESNHNSP-----IFLDRE-----TESLHYSTLFDLSLEGVPS 432
D 883 EFKNDLF--IEYSKILNDPKCRIFPKSPDIYIILDKIKQNCRESLYTLEDTESTIGILA 940
QY 433 GO-----DKVMSVYLGKICNVVACDPRVERHETLSQWRNRFSGAFAAHIGSNAPK 488
D 941 SKYFSDKKENEYLPYLEGVY-----RYETKLKLN-----YVDFSKHK 981
QY 489 QASMLLALFNGE 501
D 982 NNEKLINLFMSPE 994

RESULT 13
DPO3_STA00 STANDARD; PRT; 1436 AA.
AC Q53665; Q57110; Q9F1J9;
DT 15-DEC-1998 (Rel. 37, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE DNA POLYMERASE III POLC-TYPE (EC 2.7.7.7) (POLIII).
GN POLC.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP8;
RX MEDLINE=96084951; PubMed=7489915;
RA Pacitti D.F., Barnes M.H., Li D., Brown N.C.;
RT "Characterization and overexpression of the gene encoding
RT Staphylococcus aureus DNA polymerase III.";
RT Gene 165:51-56(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Inoue R., Kaito C., Tanabe M., Kamura K., Akimitsu N., Sekimizu K.;
RT "Genetic identification of two distinct DNA polymerases, DnaE and
RT POLC, essential for chromosomal DNA replication in Staphylococcus
RT aureus.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR REPLICATIVE DNA SYNTHESIS. THIS DNA
CC POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
CC N PYROPHOSPHATE + DNA(N).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. POLC
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D86727; BAA13160.1; -
CC EMBL; Z48003; CAAB8043.1; -
CC EMBL; AB053353; BAB20885.1; -
CC InterPro; IPR000520; Exonuclease.
CC InterPro; IPR003141; PHP_N.
CC Pfam; PF00929; Exonuclease; 1.
CC Pfam; PF02231; PHP_N; 1.
```

```
DR SMART; SM00479; EXOIII; 1.
KW SMART; SM00481; POLIIIA; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication; Hydrolase;
KW Nuclease; Exonuclease.
FT DOMAIN 420 584 EXONUCLEASE.
FT CONFLICT 207 208 NE -> QO (IN REF. 1).
FT CONFLICT 952 952 T -> K (IN REF. 1).
FT CONFLICT 1030 1049 MISSING (IN REF. 1; CAAB8043).
FT CONFLICT 1035 1035 R -> A (IN REF. 1).
FT CONFLICT 1147 1153 EFGTGFV -> NSDRIR (IN REF. 1).
FT CONFLICT 1260 1260 M -> I (IN REF. 1).
FT CONFLICT 1409 1409 S -> Y (IN REF. 1).
SQ SEQUENCE 1436 AA; 162459 MW; 2F70E034C0FBF723 CRC64;

Query Match 3.6%; Score 99.5; DB 1; Length 1436;
Best Local Similarity 20.9%; Pred. No. 37;
Matches 116; Conservative 84; Mismatches 206; Indels 149; Gaps 30;

QY 10 QDKKTMNEEDDG--NGMDLAVLGYKVRSSMADYAKLEQLEVMMSNVQEDDLSOL 67
D 416 KDATVVFVDETGLSNQYDKIIELAAVKVHGEIID--KFER---FSNPHE-RLSET 467
QY 68 ATEVHYNPAELYTWLDSMLTDLNPPSSNAEYDLKAIPGDAILNQFAIDSSASSNGGGG 127
D 468 IINLTHIT-----DDMLYDA-PEIEEVLTEFKEMWGDAL--FVAHNASFDM---G 511
QY 128 DTYTNKRL---KCSNGVET---TTATAESTRHVVLVDSOENGRLV--HALLACAEA 178
D 512 FIDTGYERLGGPSTNGVIDTLELSRTINTEYKHGLNFKLAKYGVELTQHRAIYDTEA 571
QY 179 VKENLTVAEALVKOIGFLAVSQIGAMRKVATYFAEALARIYRLSPSQSPIDH---SLS 235
D 572 TAYIFIKWVOQM--KELGVLNHNKILSNEDAYKRARPSTHTLIVQOQGLKNLFKIVS 630
QY 236 DTLQMHFYETCPYLKFAHTANOALILEAF-----QGKRHVHVIDFSMSQ----- 279
D 631 ASLVKYFYRT-PRIP-----RSLDEYREGLVGTACDEGELFTAVMQKDSQVEKIA 682
QY 280 -----GLQWALMALRPGPPVFLRTIGIGPPAPDNFDYLVHEVGCKLAHLAEA---- 329
D 683 KYDFIEIQPALYQDLIDR---ELIRD-----ETLHEIYQRLIHAGDTAGIP 728
QY 330 -----IHVEFEYRGFVANTLA-----DLDSMLELRPSIESVAVNSVFELHKLGRP 377
D 729 VIATGNAYLFEHGDGIARKILIASQPGNPLNRSTL---PEAHFRTIDMLNEFH--FLGEE 784
QY 378 GAIDKVLGVNQIKPEIFTVVVEQESNHNSPFLDRFTESLHYSTLFDLSLEGVSGODKV 437
D 785 KAHELVKNTNELADRIERVV-----PIKDELYTPRM-----EGANEE 822
QY 438 MSEV-----YLGKQICNVVACDPRVERHETLSQWRNRFSGAFAAHIGSNAPKQA 490
D 823 IRELSYANARKLYGEDLPQIVI---DELEK-----ELKSIIGN-GFAVIVLISORLVKK 872
QY 491 SMLLALFNGEGRYV 505
D 873 SL-----DDGYLV 880

RESULT 14
SPYA_FELCA STANDARD; PRT; 414 AA.
AC P41689;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SERINE--PYRUVATE AMINOTRANSFERASE, MITOCHONDRIAL PRECURSOR
DE (EC 2.6.1.51) (SPT) (ALANINE--GLYOXYLATE AMINOTRANSFERASE)
DE (EC 2.6.1.44) (AGT).
GN AGXT OR AGT1.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
NCBI_TaxID=9685;
[1]
SEQUENCE FROM N.A.
TISSUE=Liver;
Lund M.J., Purdue P.E., Danpure C.J.;
"Molecular evolution of alanine/glyoxylate aminotransferase 1
Intracellular targeting. Analysis of the feline gene.";
Eur. J. Biochem. 221:53-62(1994).
-f- FUNCTION: DUAL METABOLIC ROLES OF GLUCONEOGENESIS (IN THE
MITOCHONDRIA) AND GLYOXYLATE DETOXIFICATION (IN THE PEROXISOMES).
-l- CATALYTIC ACTIVITY: L-SERINE + PYRUVATE = 3-HYDROXYPYRUVATE +
L-ALANINE.
-l- CATALYTIC ACTIVITY: L-ALANINE + GLYOXYLATE = PYRUVATE + GLYCINE.
-l- COFACTOR: PYRIDOXAL PHOSPHATE.
-l- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (>90%) AND PEROXISOMAL.
-l- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT
AMINOTRANSFERASES.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL: X75923; CAA53527.1; "
InterPro: IPR000192; Aminotransf_class_v.
Pfam: PF00266; aminotran_5; 1.
PROSITE; PS00342; MICROBODIES_CTER; 1.
PROSITE; PS00595; AA_TRANSFER_CLASS_5; 1.
KW Transferrase; Aminotransferase; Pyridoxal phosphate; Pteroisomase;
KW Mitochondrion; Transit peptide; Alternative initiation.
FT TRANSIT 1 23 MITOCHONDRION (BY SIMILARITY)..
FT CHAIN 24 414 SERINE--PYRUVATE AMINOTRANSFERASE,
FT FT MITOCHONDRIAL ISOFORM.
FT CHAIN 23 414 SERINE--PYRUVATE AMINOTRANSFERASE,
FT FT PEROXISOMAL ISOFORM.
FT INIT_MET 23 23 FOR PEROXISOMAL ISOFORM.
FT BINDING 231 231 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT SITE 412 414 MICROBODY TARGETING SIGNAL (BY
FT FT SIMILARITY).
SQ SEQUENCE 414 AA; 45507 MW; OD1B01E0EA199B3 CRC64;
Query Match 3.6%; Score 99; DB 1; Length 414;
Best Local Similarity 21.9%; Pred. No. 6.6;
Matches 70; Conservative 43; Mismatches 113; Indels 94; Gaps 15;
QY 292 LRPGPPVFRLTGIPGP--AP-----DNFDYLHEVGCKLAHLAEAIHVEPEYRGFVA 341
: : : : : : : : : : :
Db 37 LRPLSTPNRLLLGGPSNLAPRVLVAGGKMGIMHMK---EMFOIMDDIKOGIYYVFQTK 93
: : : : : : : : : : : : : : : : :
QY 342 NTL-----ADLDASMLE-LRPSEIESVAVNSVP-----ELHKLLGRGP 378
: : : : : : : : : : : : : :
Db 94 NPLTLAISGSGHCALAEALFNILEPGLDPFLGVNGINGQRAADIGERIGARVHPMIKDPG 153
: : : : : : : : : : : : : : : : :
QY 379 ---AIDKVLGVNQIKPEITVVQECSNNHSPIFLDFRTSLHYISTLF-----DSLEGV 430
: : : : : : : : : : : : : : : : :
Db 154 NHYTLQLEELAQAHRPVLLFTQGESSGVLPQLDGYGELCHRYNCLLLVDVSVASLCGT 213
: : : : : : : : : : : : : : : : :
QY 431 P-----SGDGKVM-----EVILG-KQICNVV 451
: : : : : : : : : : : : : :
Db 214 PIYWDOGDIVLYSGSGCVLNLSPPGTSLSFSDKANKKIYTRKTVPVSFYDLDMKLANII 273
: : : : : : : : : : : : : : : : :
QY 452 ACDGPORVERHET--LSQWRNRFGSAGFAAAHIGSNAFKAQSMALLALFNG---GEGYRVE 506
: : : : : : : : : : : : : : : : :
Db 274 GCDGKPRIYHTTPVSVLSYLSRESLALIAEQGL-ENSWRQHREVTALHGRLQGLGLQLF 332
: : : : : : : : : : : : : : : : :
QY 507 ESDGCLMLGWHTHTRIATSA 526
: : : : : : : : : : : : : :

[illegible]

Db 451 LIWPFVNSYLK-----GQEPDPDLLFWNADSTRMAAANHAFYLRNCYLRNA-LTON 502
QY 258 QAILEAFQGGK-----RVHVIDEFSMSOGLQWPALMQALALR-PGGPPVFRLTGIGPPA- 309
Db 503 EMILD---GKRISLKDVKIPYINLATREDHIAPAKSVFLGSRFFGGKVEFVVTGSGHIAG 559
QY 310 ----PDNFDY 315
Db 560 VVNPPDKRKY 569

Search completed: May 8, 2002, 11:11:55
Job time: 215 sec